2020 Census Detailed Operational Plan for: 13. Non-ID Processing Operation (NID)

A New Design for the 21st Century

Issued: August 24, 2016
Version: FINAL v1.0
Prepared by: Decennial Census Management Division

United States™ Census Bureau

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

United States™ Census 2020

[Page intentionally left blank.]

Approvals

This NID Detailed Operational Plan has been reviewed and approved for use.

Francis C. McPhillips	August 1, 2016
Francis C. McPhillips IPT Lead	Date Signed

K. Evan Moffett	August 3, 2016
K. Evan Moffett IPT Program Manager	Date Signed

Deborah M. Stempowski	August 5, 2016
Deborah M. Stempowski, Chief Decennial Census Management Division	Date Signed

Lisa M. Blumerman	August 7, 2016
Lisa M. Blumerman Associate Director for Decennial Census Programs Chair, Portfolio Management Governance Board	Date Signed

Document Change History

Revision #	Version	Date	Description
1	v0.01c	May 3, 2016	Initial Working DRAFT Version from 2020 Census DOP template. Uses Annotated version of NID BPM Version 5.0, dated 4/25/2016.
2	v0.02b	May 23, 2016	Initial Review DRAFT Version from 2020 Census DOP template. Uses Annotated version of NID BPM Version 5.0, dated 4/25/2016.
3	v0.03b	July 6, 2016	Updated DRAFT Version from 2020 Census DOP template. – incl. minor graphics updates and resolution of TBD items. Uses Annotated version of NID BPM Version 5.0, dated 4/25/2016.
4	V0.04b	August 23, 2016	Updates based on feedback from senior leadership. DRAFT Version form 2020 Census DOP template. – incl. minor updates to content.
5	V1.0	August 24, 2016	Final Version

Note: Edit the fields below to update the Document Version, Date and Status in the Page Footers throughout the document.

DocVersion:	Version v1.0 (Final)
DocDate:	August 24, 2016
DocStatus:	Final

Table of Contents

List of Figures

List of Tables

[Page intentionally left blank.]

1. Document Purpose

The 2020 Census Detailed Operational Plan for the Non-ID Processing Operation (NID) is intended for use by U.S. Census Bureau managers, staff, contractors, and other internal and external stakeholders working on the 2020 Census. The document presents the detailed operational design for the 2020 Census NID Operation and includes a summary of the operational processes involved, their inputs, outputs and controls, and the basic mechanisms employed to conduct the operational work.

Anticipated uses of this document include the following:

- Communication – Documents operational design details for internal and external stakeholders
- Planning – Documents planning assumptions and key milestones
- Staffing – Documents staffing needs and strategies
- Design – Describes operations and flows, which inform design of IT systems, manual processes, and training
- Development – Identifies business rules and required capabilities to be developed
- Testing – Provides a basis for developing integrated test plans for IT systems and processes

This document complements the 2020 Census Operational Plan, which presents the initial baseline version of the 2020 Census operational design and covers all operations required to execute the 2020 Census, starting with precensus address and geographic feature updates and ending once census data products are disseminated and coverage and quality are measured.

This document will be updated over time to reflect changes in strategies that result from 2020 Census planning, research, and testing activities.

2. Operational Overview

2.1 Operation Purpose

The Non-ID Processing operation is focused on making it easy for people to respond anytime, anywhere to increase self-response rates. The operation enables the 2020 Census mission by:

- Providing response options that do not require a unique Census identification (ID) code
- Maximizing real-time matching of Non-ID respondent addresses to the Census Master Address File (MAF) and assigning to census blocks
- Accurately assigning nonmatching addresses to census blocks
- Conducting validation of all Non-ID responses

The 2020 Census respondents will have the opportunity to fill out their census questionnaire online using their preassigned Census ID. In addition, respondents will be given the option of going to the Census website and completing the questionnaire without an ID. These cases without an ID are commonly referred to as "Non-ID Cases."

The process of comparing these Non-ID Cases to address records contained in the Master Address File (MAF)/Topologically Integrated Geographic Encoding and Referencing (TIGER) database (MTdb) to determine whether they match existing address records and/or assigning census geographic codes to these cases is known as "Non-ID Processing." While this option provides respondents with more flexibility and ease in responding, it could also increase the Non-ID workload over 2010 Census levels.

The U.S. Census Bureau's 2020 Census Operational Plan documents the current design for conducting the 2020 Census. As the initial version of an emerging concept of operations, it reflects and supports evidence-based decision making by describing design concepts and their rationale, identifying decisions still to be made, and describing significant issues and risks related to the implementation of the Operational Plan.

Accordingly, the Non-ID Processing operation's goal is to use real-time automated processing that will readily link (match) Non-ID addresses to an existing MTdb record, or in the case of a non-match, otherwise determine the correct physical location and associated geographic codes. In addition, the operation will implement tools and techniques for performing interactive matching and geocoding when automated processing is unable to resolve a Non-ID case. The

operation also leverages office-based solutions to reduce the field verification[1] workload for the nonmatched/geocoded Non-ID cases from automated and interactive matching and geocoding. Lastly, the operation shall attempt to ensure that the increase in the proportion of Non-ID response does not expose the census to increased fraud.

2.2 Background

During the 2010 Decennial Census enumeration for stateside (the 50 states and the District of Columbia) and Puerto Rico, as well as during previous censuses, the majority of people self-responded by questionnaires delivered to their living quarters (LQ) by mail or hand-delivered by U.S. Census Bureau field staff. These questionnaires contained a Census ID that linked the questionnaire to the address of the LQ. While the vast majority of responses in 2010 fell into the category of responses containing a Census ID, there were a number of individuals who responded via methods (such as submitting a "Be Counted" form) resulting in a response lacking a preassigned Census ID.

The ultimate goal of Non-ID Processing is to determine if the addresses submitted by respondents could be associated with an existing MTdb record in the Census universe, or if not matched to a MTdb record but subsequently verified, assigned a new Census ID, geocoded and added to the census universe. It is critical the addresses are associated with geographic codes in the correct physical location because the census data product relies on the accuracy of these geographic codes.

2.3 Design Overview

The sections below present the high-level design for the Non-ID Processing Operation (NID). Please refer to the 2020 Census Operational Plan for a complete design for all 2020 Census operations.

2.3.1 High-Level Operational Design

The design of the NID Operation for the 2020 Census includes three major operational activity areas:

- Real-Time Non-ID Processing Phase

[1] Field verification existed in the 2000 and 2010 Censuses to account for Non-ID responses that were unable to be matched to the MAF, or matched to a previously ungeocoded record. In these instances, this was the mechanism to ensure the respondent's address/living quarters actually exist and are physically located and associated with the correct geographic codes assigned to them.

- Post Real-Time Non-ID Processing Phase
- Non-ID Response Validation Phase

Each of these major activity areas is summarized below. Together, the activities in these three areas represent the complete set of work that needs to be performed to conduct this operation.

Real-Time Non-ID Processing Phase

The Real-Time Non-ID Processing (RTNP) phase will attempt to resolve Non-ID cases submitted by the respondent using matching and geocoding services at the time of submission. This phase also enables a feedback loop to the respondent which is utilized to increase the rate of successful matching and geocoding during the response.

Post Real-Time Non-ID Processing Phase

The Post Real-Time Non-ID Processing phase will attempt to resolve Non-ID cases submitted by the respondent that are not resolved in the Real-Time Non-ID Processing phase. This phase utilizes what is referred to as "Asynchronous Non-ID Processing" that is the same matching and geocoding services as Real-Time but also includes the use of a composite of federal administrative records and third-party data to obtain corrected or missing address data in order to improve the match/geocode rate.

This phase also uses what is referred to as "Clerical Non-ID Processing" that attempts to obtain an address match and/or geocode for Non-ID address records not resolved during Automated Non-ID Processing (both Real-Time and Asynchronous). This work will be achieved using available geographic reference materials and an interactive matching and geocoding application in an office-based operation.

Non-ID Response Validation Phase

The Non-ID Response Validation phase will attempt to determine whether fraudulent returns have been submitted during the 2020 Census. This phase will use existing processes and expand processes where needed, establishing criteria, thresholds, and level of investigation that may be needed. The investigation may include the possibility for a field follow-up to review for potential fraud.

The full hierarchy of activities for the NID Operation is provided in Appendix C in the form of an Activity Tree. In the Activity Tree, each major operational activity listed above is numbered and then decomposed into a numbered set of subactivities, some of which are further decomposed into more detailed numbered subactivities or steps.

For a full description of the operational subactivities that comprise the NID Operation, see the Detailed Process Description discussions in Section 3 below.

2.3.2 NID Operational Context

The NID operational activities described above are conducted within the context of other 2020 Census operations and other programs or data sources that are external to the 2020 Census Program. One way to depict an operational context is by using a "Context Diagram," which shows the boundary of the operational process, the operational activities it contains, and the information exchanged with its neighbor operations (or other entities) as well as the resources (mechanisms) needed to conduct the operational work.

Figure 1 is a top-level context diagram for the NID Operation represented as an Integrated Definition, Level 0 (IDEF0) model. An IDEF0 Model of a process (or operation) shows the Inputs, Controls, Outputs and Mechanisms of the process. These IDEF0 model elements are summarized below and described further in the sections that follow.

The yellow box in the center of the IDEF0 model lists the major operational activity areas for the operation, numbered as given in the NID Operation Activity Tree in Appendix C. Specific Information Exchanges (IE) are shown in different colored boxes to represent the Inputs (green boxes on left side), Outputs (orange boxes on right side), Controls (purple boxes on top) and Mechanisms (blue boxes on the bottom). Boxes to the left of the Inputs indicate the *Provider* of the inputs to the operation (typically another 2020 Census operation or an external source). The Provider of the Controls is noted in the box itself. Boxes to the right of the Outputs indicate the *Receiver* of the outputs (typically another 2020 Census operation or external entity). Each Information Exchange has a name and a unique number for identification purposes.

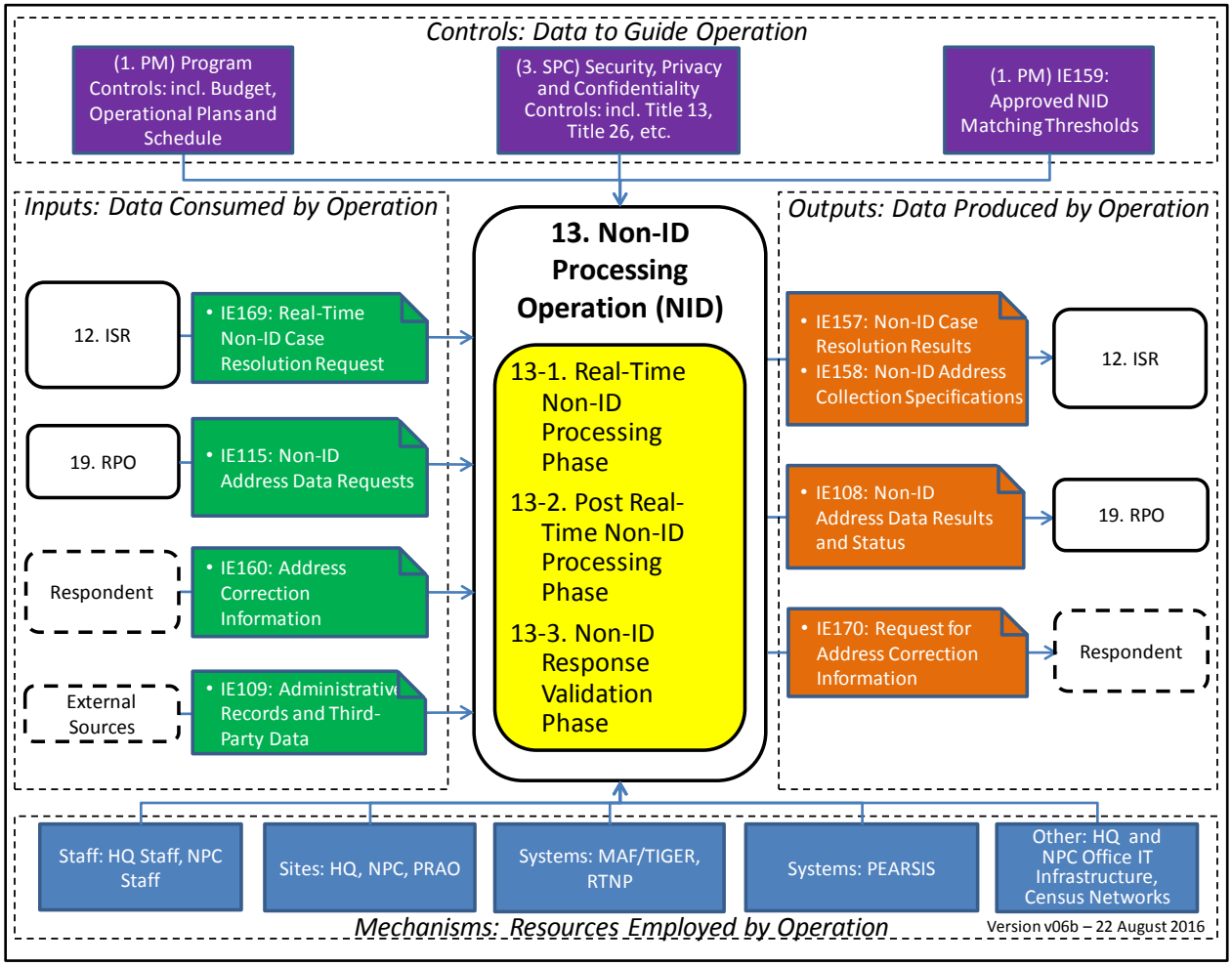

Figure 1: Non-ID Processing Operation (NID) Context Diagram

For NID, the input data consists of Real-Time Non-ID case resolution requests from Internet Self Response (ISR), Non-ID address data requests from Response Processing Operations (RPO), and address correction information from the respondent as well as administrative records and third-party data for use in Non-ID related matching activities.

The control data which is used to guide the NID activities includes Program Controls and Security Controls as well as approved matching thresholds.

The output data from NID consists of Non-ID case resolution results and Non-ID address collection specifications to ISR, Non-ID address data results and status to RPO, and request for address correction information to the respondent.

The mechanisms (i.e. physical resources) employed during Non-ID Processing include staff (at HQ and NPC), infrastructure sites (HQ and NPC), and Systems (MAF/TIGER, RTNP and

Production Environment for Administrative Record Staging, Integration and Storage (PEARSIS). Other mechanisms employed include HQ, NPC, and PR Area Office IT infrastructure and Census Bureau networks.

For detailed descriptions of the Inputs, Controls, Outputs and Mechanisms used by the NID Operation, see the sections that follow.

2.3.2.1 NID Operational Inputs

Inputs are the data that are consumed by the operation. The inputs define the amount of operational work that needs to be performed.

Table 1 lists the inputs to the NID Operation.

Table 1: NID Operational Inputs

Provider	Information Exchange	Description
12. Internet Self Response Operation (ISR)	IE169: Real-Time Non-ID Case Resolution Request	Non-ID address from a Respondent or from a Census Questionnaire Assistance (CQA) Agent using an ISR application.
19. Response Processing Operation (RPO)	IE115: Non-ID Address Data Requests	Non-ID address from Response Processing.
Respondent	IE160: Address Correction Information	Corrected address from respondent.
External Sources	IE109: Administrative Records and Third-Party Data	Administrative Records data to be used by PEARSIS for Non-ID related matching activities.

2.3.2.2 NID Operational Controls

Controls are the data that guide the behavior of the operation. They are not consumed by the operation, but rather they provide guidance, models, limits, criteria, cutoff dates, or other information that controls the way in which the operational work is performed.

Table 2 lists the controls for the NID Operation.

Table 2: NID Operational Controls

Provider	Information Exchange	Description
1. Program Management Operation (PM)	Program Controls	Program Control information including: • Budgets • Operational Plans and Schedule • Quality Measures
3. Security, Privacy, and Confidentiality Operation (SPC)	Security, Privacy and Confidentiality Controls	Laws, policies, regulations, and guidelines related to physical security, IT security, data security and privacy and confidentiality impacts, analyses, and processes. These include but are not limited to Title 13, Title 26, and other laws and policies related to protection of personally identifiable information.
1. Program Management Operation (PM)	IE159: Approved NID Matching Thresholds	Approval of NID Operation Matching Thresholds as provided by PM.

2.3.2.3 NID Operational Outputs

Outputs are the data produced by the operation. The outputs constitute the results of operational work that has been performed. Outputs produced may be used as inputs or controls to other operations.

Table 3 lists the outputs from the NID Operation.

Table 3: NID Operational Outputs

Consumer	Information Exchange	Description
12. Internet Self Response Operation (ISR)	IE157: Non-ID Case Resolution Results	Non-ID matching and geocoding results from RTNP.
	IE158: Non-ID Address Collection Specifications	Non-ID address collection specifications for internet instrument.

Consumer	Information Exchange	Description
19. Response Processing Operation (RPO)	IE108: Non-ID Address Data Results and Status	Non-ID matching and geocoding results from post real-time Non-ID processing.
Respondent	IE170: Request for Address Correction Information	Request for feedback from respondent on completeness/accuracy of their submitted address.

2.3.2.4 NID Operational Mechanisms

Mechanisms are the resources (people, places and things) that are used to perform the operational processes. They include Staff Resources, Infrastructure Sites, and Systems and other Technology Infrastructure.

Staff Resources

Table 4 identifies the Staff Resources employed for the NID Operation.

Table 4: Staff Resources used within NID Operational Activities

Staff Resources	Description/Role
HQ Staff	HQ Staff to manage overall NID Operation and coordinate activities with NPC for the Clerical Operation. HQ Staff to conduct analysis work.
NPC Staff	NPC Staff to manage and conduct NID Operation clerical work.

Infrastructure Sites

Table 5 identifies the Infrastructure Sites employed for the NID Operation.

Table 5: Infrastructure Sites for NID Operational Activities

Infrastructure Site	Description/Role
HQ	HQ Sites for office work
NPC	National Processing Center site used for clerical work
Puerto Rico Area Office (TBD)	Possible use of Census field office space in Puerto Rico Area Office (PRAO) for clerical work specific to Non-ID cases from PR

Systems and other Technology Infrastructure

Table 6 identifies the Systems employed for the NID Operation.

Table 6: Systems used within NID Operational Activities

System	Description
MAF/TIGER: Master Address File/ Topologically Integrated Geographic Encoding and Referencing System	The MAF/TIGER System provides the corporate address list, the map data, the geocoding service, and the distribution of related geographic and address products.
RTNP: Real-Time Non-ID Processing	Real-Time Non-ID Processing provides resolution of Non-ID responses at the time of collection (during ISR).
PEARSIS: Production Environment for Administrative Records Staging, Integration and Storage	PEARSIS provides a repository of federal administrative records and third party data as well as business logic for comparing that data with respondent-provided data, such as address, respondent name, telephone number, etc. PEARSIS serves two purposes for Non-ID: it can provide corrected address information to enhance the respondent-provided data, and it can also serve as one of the mechanisms to validate census responses which lack an ID.

Other Technology Infrastructure employed for the NID Operation includes:

- HQ and NPC Office IT Infrastructure for conducting NID operational work
- Census Network connectivity for data transmission between operational systems and operational sites

2.4 NID Data Flow and Operational Influences

Note to Reader: An Integrated Operations Diagram (IOD) for the data collection operations is being developed as part of the Response Processing Operation (RPO) DOP, which will not be completed until fiscal year 2017. This diagram, which will show the flow of information among all of the data collection operations, is intended to help the reader understand how this operation fits into the bigger picture. The Data Collection IOD will be added as Figure 2 in the next release of this document.

<Data Collection IOD to be inserted here at a later date>

Figure 2: 2020 Census Data Collection – Integrated Operations Diagram (IOD)

2.5 NID Design Assumptions and Constraints

2.5.1 Assumptions

- 2020 Non-ID Processing will account for non-ID response from these modes including, but not limited to the following:
    - Internet
    - Respondents who provide a Census ID (e.g., from a Census Bureau-provided postcard) but do not confirm the address associated with that ID when prompted by the Internet instrument. Instead, they provide a different living quarters address for their household, thus "spawning" a Non-ID case from an ID'ed response.
    - Telephone interviews originating from Census Questionnaire Assistance calls
    - Respondents who do not provide a Census ID when prompted by the internet instrument (including respondents calling Census Questionnaire Assistance (CQA).
    - Paper questionnaires received at data capture centers that lack a Census ID (e.g., exception check-in cases where the barcode was damaged, missing, or is otherwise unreadable).
    - Living quarters added to the universe by enumerators during fieldwork.
- Mechanisms for real-time matching and geocoding will be developed or are readily available, including the following:
    - Real-time address standardization
    - Real-time access to MTdb data for matching and geocoding Non-ID cases
- Applications and infrastructure associated with Non-ID Processing will safeguard Title 13 information.
- Administrative record data of sufficient quality and quantity will be available to Automated Non-ID Processing to enable the process of correcting erroneous information in some of the respondent-provided addresses, as well as supply missing data items.
- Non-ID cases not resolved (i.e., matched to a MAF record with a confirmed census block geocode) during automated processing will be passed to a manual Non-ID processing operation, which requires Census Bureau staff to interactively match and geocode addresses from Non-ID responses.

2.5.2 Constraints

- Census test sites over the course of the decade cannot completely represent the full set of challenges Non-ID processing will encounter on a national scale during the 2020 Census. For example, there are a number of anomalous address systems used in specific geographic areas (e.g., hyphenated address numbers in Queens, New York, or the quadrant system used in many parts of Utah). However, the sites selected so far, as well

as the criteria discussed for the remaining tests should provide an opportunity to examine many of the challenges anticipated, such as in Puerto Rico and American Indian reservations.

- Administrative records coverage will vary by geographic area as well as source. Therefore, the ability to correct or supplement respondent address data will be subject to the limitations of coverage.
- Use of administrative records in Non-ID processing will be governed by constraints associated with the sources of data.
- Use of location services from mobile devices, or use of online services to derive a coordinate location for an address will be governed by existing policy and legal constraints.

3. Non-ID Processing Operation (NID) Detailed Process Description

Figure 3 is a top-level Business Process Model (BPM) showing the Level 1 activity areas within the NID Operation. BPMs for the 2020 Census follow industry-standard Business Process Model and Notation (BPMN). Refer to Appendix D for an explanation of how to read the BPMN notations and a copy of all of the BPMN diagrams for this operation.

This top-level BPM serves as the Context Model for the NID Operation. A BPMN Context Model displays the high-level activities within the operation and relationships between them, whereas the IDEF0 Context Diagram shown earlier depicts the boundaries of the operation or activity and the interfaces between the operation or activity and other operations and activities with which it is associated.

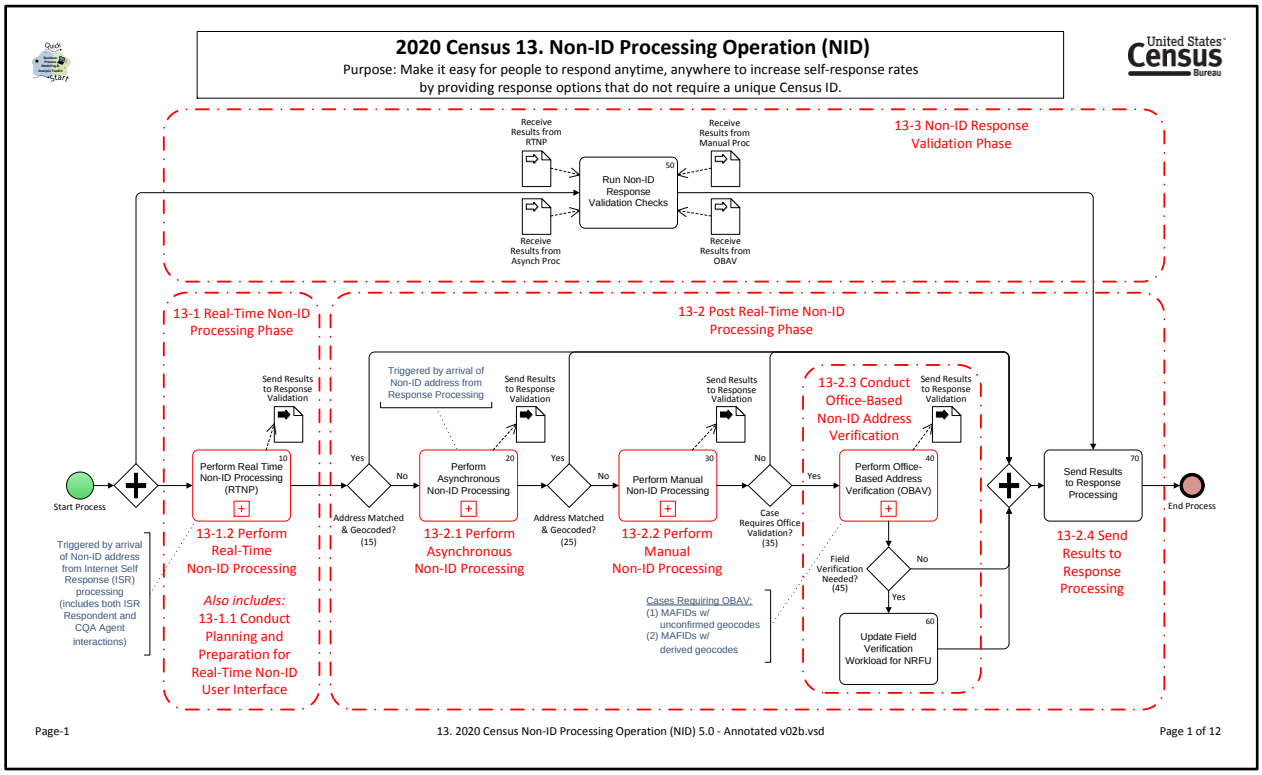

Figure 3: NID Operation Context Model

The NID Operation is subdivided into the following Activity Areas.

- Real-Time Non-ID Processing Phase [13-1]
- Post Real-Time Non-ID Processing Phase [13-2]

- Non-ID Response Validation Phase [13-3]

The business processes for each of these Level 1 activity areas are discussed along with their inputs and outputs in the following sub-sections.

3.1 Real-Time Non-ID Processing Phase [13-1]

Figure 4 shows the BPM for the Real-Time Non-ID Processing Phase [13-1] activity area (within the shaded Gray Rounded Rectangle) and its constituent activities within the overall context of the NID Operation.

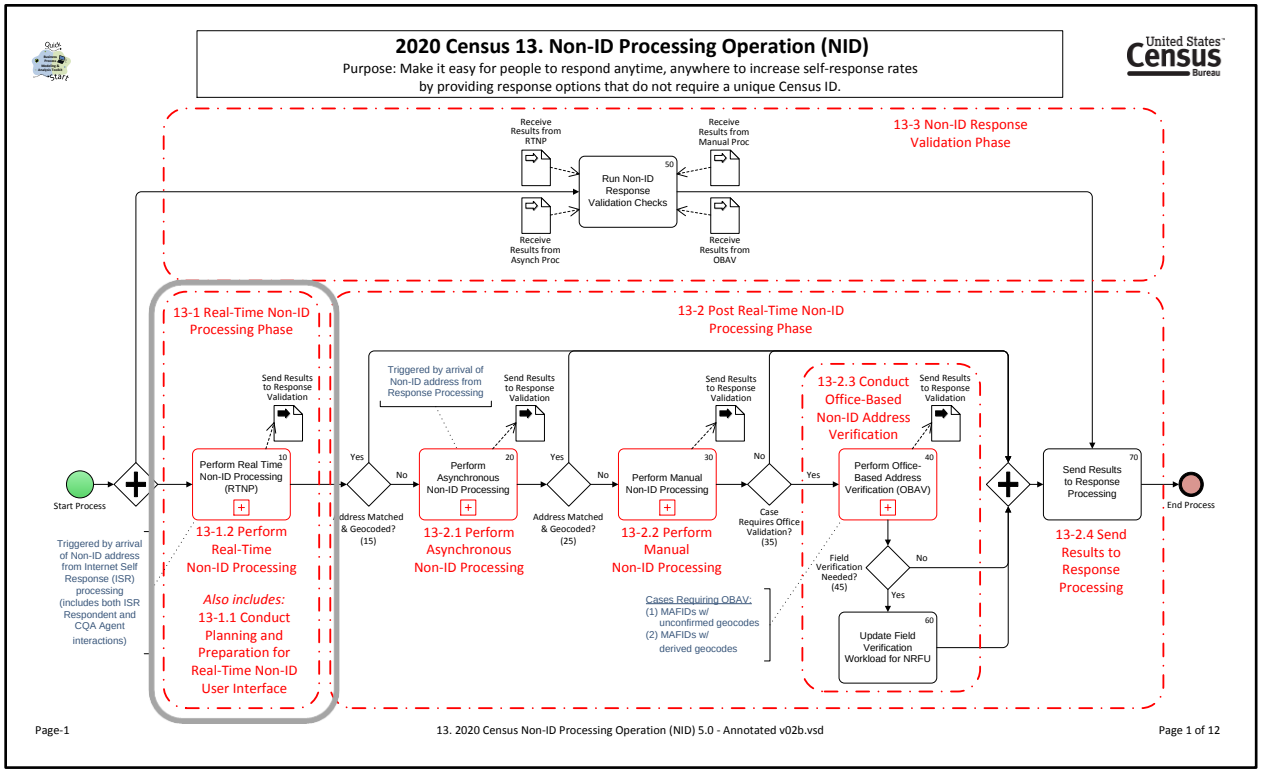

Figure 4: Real-Time Non-ID Processing Phase [13-1] Constituent Activities

The Real-Time Non-ID Processing (RTNP) Phase is subdivided into the following Activity Areas.

- Real-Time Non-ID Processing Phase [13-1]
    - Conduct Planning and Preparation for Real-Time Non-ID User Interface [13-1.1]
    - Perform Real-Time Non-ID Processing (RTNP) [13-1.2]
        - Perform Real-Time Address Standardization [13-1.2.1]
        - Perform Real-Time Non-ID Matching and Geocoding [13-1.2.2]

Subsequent sections describe the Real-Time Non-ID Processing operational sub-activities in detail.

3.1.1 Conduct Planning and Preparation for Real-Time Non-ID User Interface [13-1.1]

NID Operation staff works with ISR on the development of specifications for RTNP address data collection.

3.1.2 Perform Real-Time Non-ID Processing (RTNP) [13-1.2]

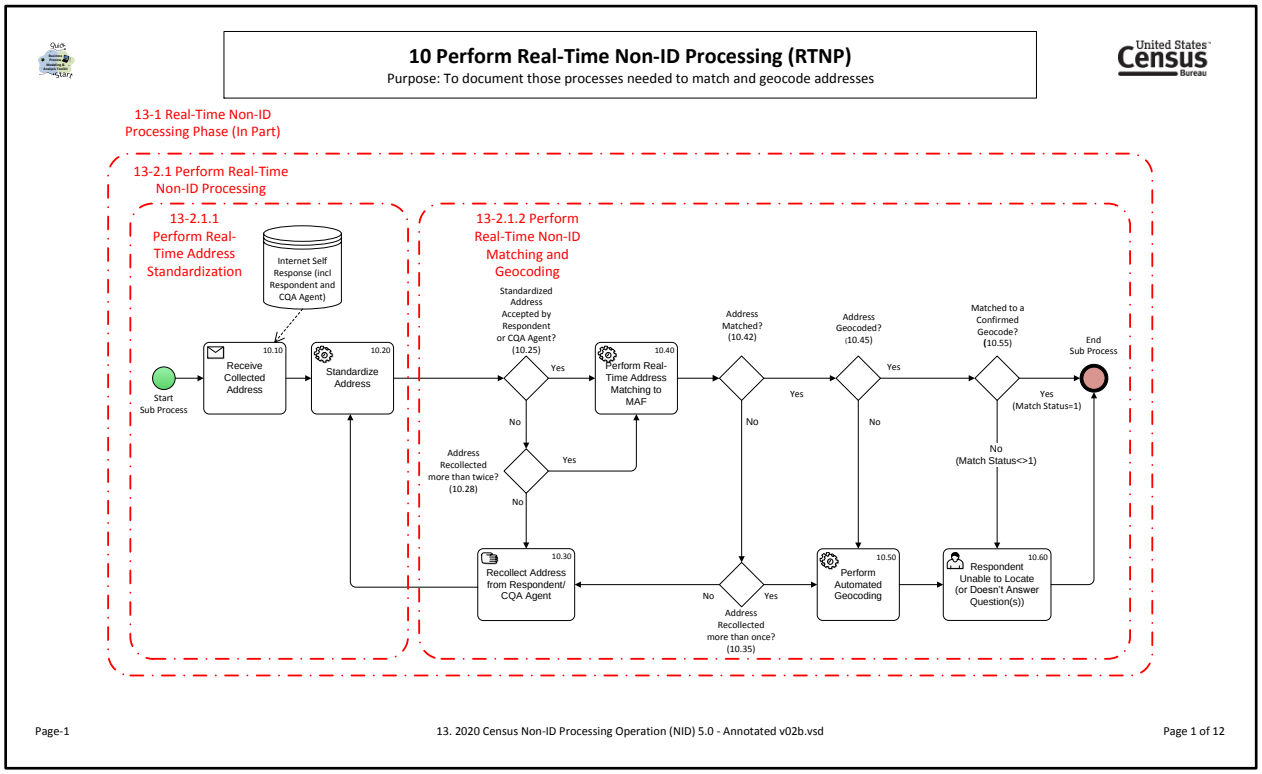

Figure 5: Perform Real-Time Non-ID Processing (RTNP)

Perform Real-Time Non-ID Processing (RTNP) is subdivided into the following Activity Areas.

- Perform Real-Time Non-ID Processing (RTNP) [13-1.2]
  - Perform Real-Time Address Standardization [13-1.2.1]
  - Perform Real-Time Non-ID Matching and Geocoding [13-1.2.2]

3.1.2.1 Perform Real-Time Address Standardization [13-1.2.1]

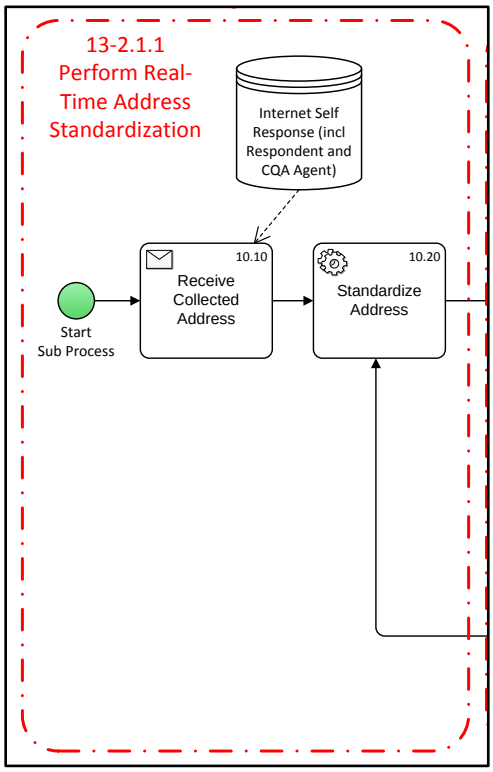

Figure 6: Perform Real-Time Address Standardization

As shown in the BPM above, there are two steps involved in this activity

- Receive Collected Address [10.10]
- Standardize Address [10.20]

Real-Time Address Standardization receives collected addresses from ISR (which includes responses directly from respondents and CQA agents). Addresses are put into a standard format (proper abbreviations, placement of address components, and layout) in preparation for matching and geocoding. This standard format provides the consistency required to match to the MAF/TIGER database.

3.1.2.2 Perform Real-Time Non-ID Matching and Geocoding [13-1.2.2]

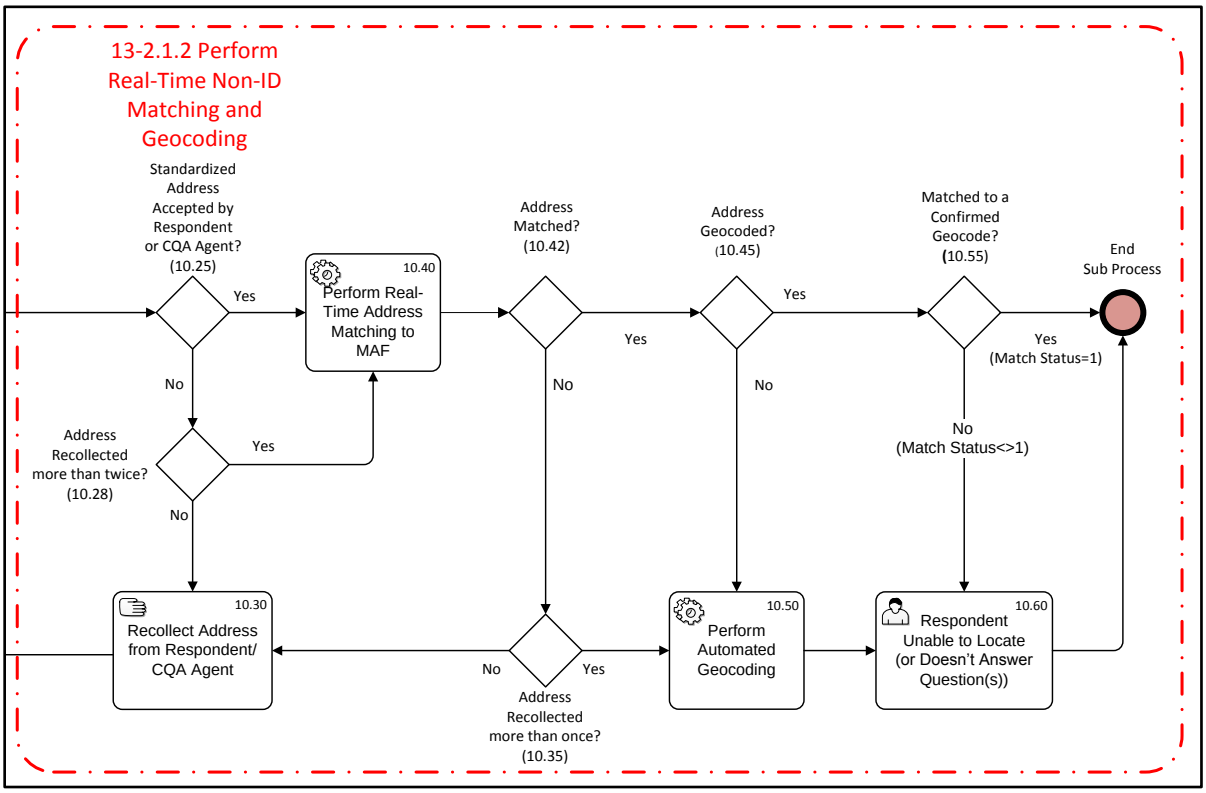

Figure 7: Perform Real-Time Non-ID Matching and Geocoding

As shown in the BPM above, there are multiple steps involved in this activity

- Recollect Address from Respondent/CQA agent [10.30]
- Perform Real-time Address Matching to MAF [10.40]
- Perform Automated Geocoding [10.50]
- Respondent Unable to Locate (or Doesn't Answer Question(s)) [10.60]

If the respondent or CQA agent does not accept the standardized address, the multifunctional system will collect a new/updated address from the respondent or CQA agent no more than two times. Once either the respondent or CQA agent accepts the standardized address or the respondent reaches the collection threshold, there is an attempt to match the address to the MAF in real-time. If the address did not match a record in the MAF and the respondent has not reached the collection threshold and then provides a new/updated address, there could be another attempt to match.

If the address is matched to a record in the MAF with a confirmed geocode, the matching and geocoding process is complete. If the address matches a record in the MAF but this record does

not have a geocode in MTdb, we will attempt to geocode the address in real-time. If the address was not geocoded in real-time or matched to a record in the MAF that had an unconfirmed geocode, either the respondent or CQA agent attempts to geocode the address in real-time using the interactive map interface built into the response instrument. If the respondent or CQA agent is unable to locate the address or does not attempt to geocode the matching and geocoding process is complete. Any non-ID respondent address which does not match to a MAF record during real-time will be passed to the next phase: Post Real-Time Non-ID Processing.

3.2 Post Real-Time Non-ID Processing Phase [13-2]

Figure 8 shows the BPM for the Post Real-Time Non-ID Processing Phase [13-2] activity area (within the shaded gray rounded rectangle) and its constituent activities within the overall context of the NID Operation.

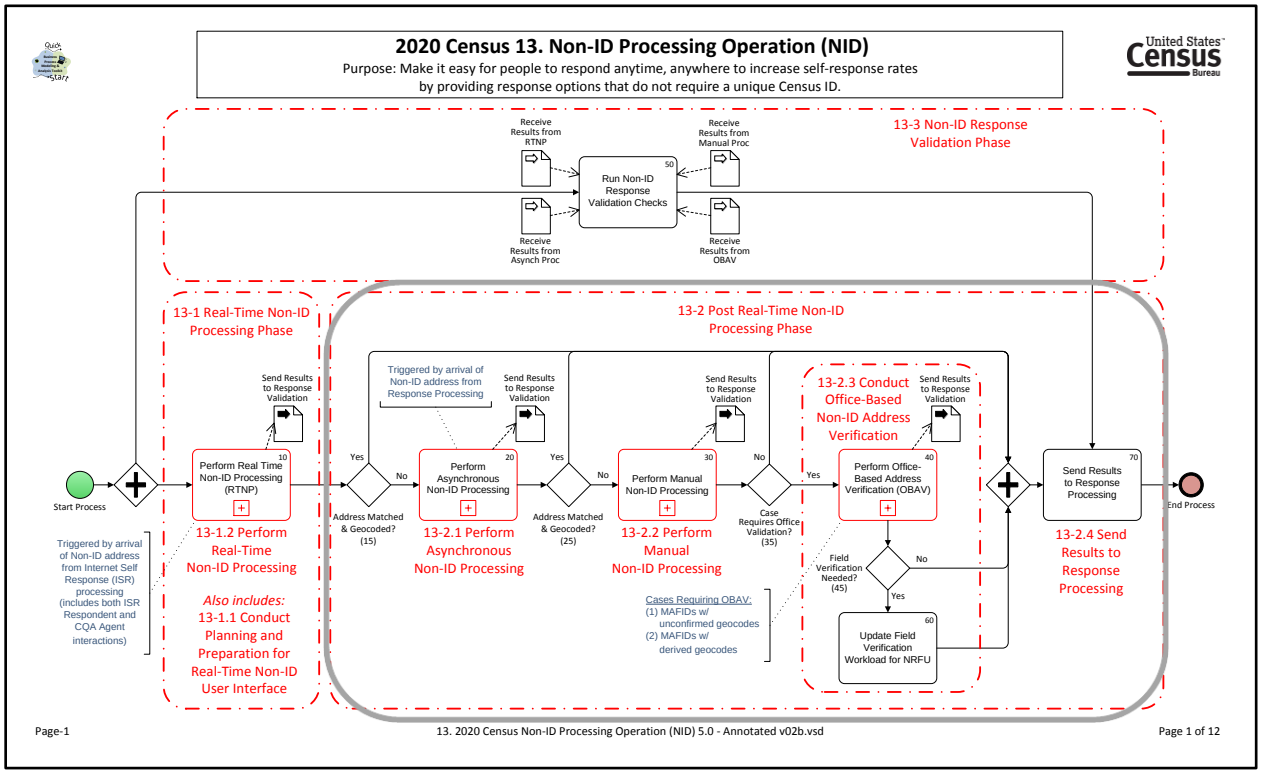

Figure 8: Post Real-Time Non-ID Processing Phase [13-2] Constituent Activities

The Post Real-Time Non-ID Processing Phase is subdivided into the following Activity Areas.

- Post Real-Time Non-ID Processing Phase [13-2]
    - Perform Asynchronous Non-ID Processing [13-2.1]
    - Perform Manual Non-ID Processing [13-2.2]
    - Conduct Office-Based Non-ID Address Verification [13-2.3]

- Send Results to Response Processing [13-2.4]

Subsequent sections describe the Post Real-Time Non-ID Processing operational subactivities in detail.

3.2.1 Perform Asynchronous Non-ID Processing [13-2.1]

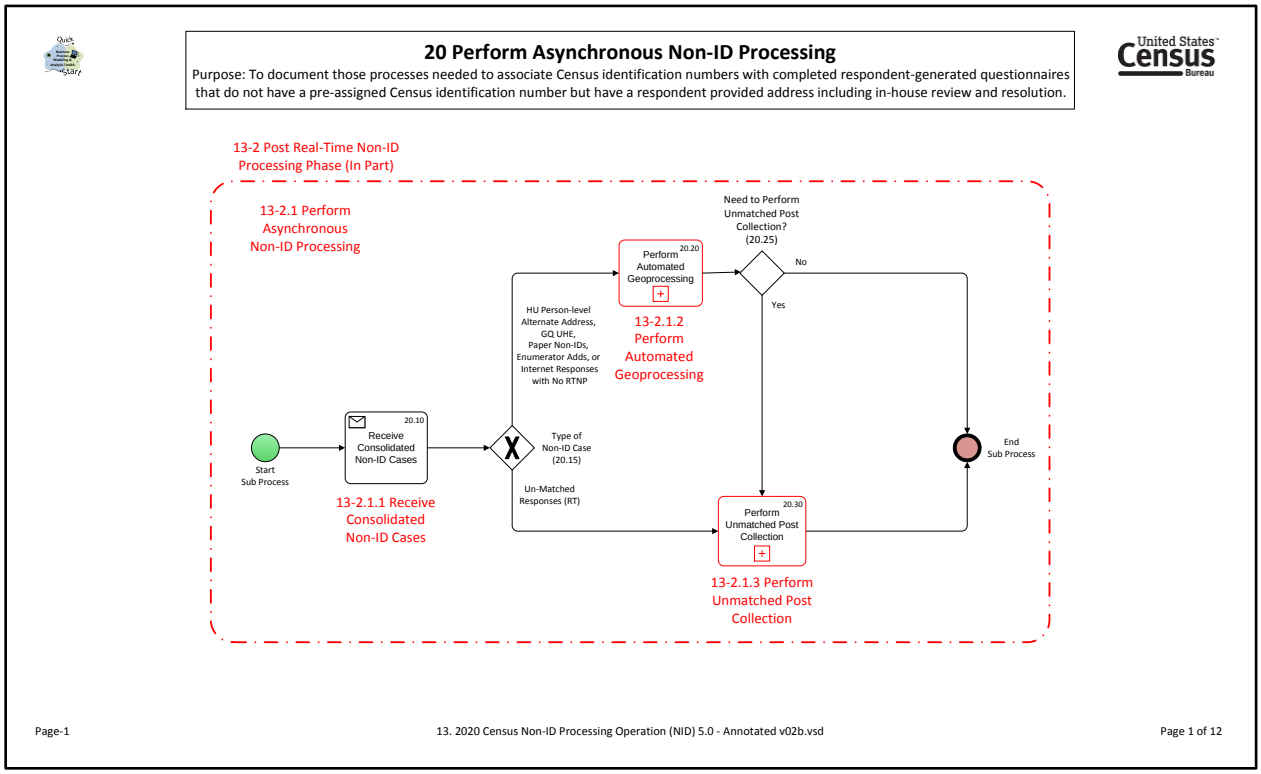

Figure 9: Perform Asynchronous Non-ID Processing

Perform Asynchronous Non-ID Processing is subdivided into the following Activity Areas.

- Perform Asynchronous Non-ID Processing [13-2.1]
  - Receive Consolidated Non-ID Cases [13-2.1.1]
  - Perform Automated Geoprocessing [13-2.1.2]
  - Perform Unmatched Post Collection [13-2.1.3]

3.2.1.1 Receive Consolidated Non-ID Cases [13-2.1.1]

As shown in the BPM there is one step involved in this activity:

- Receive Consolidated Non-ID Cases [20.10]

Asynchronous Non-ID Processing receives addresses from Response Processing.

3.2.1.2 Perform Automated Geoprocessing [13-2.1.2]

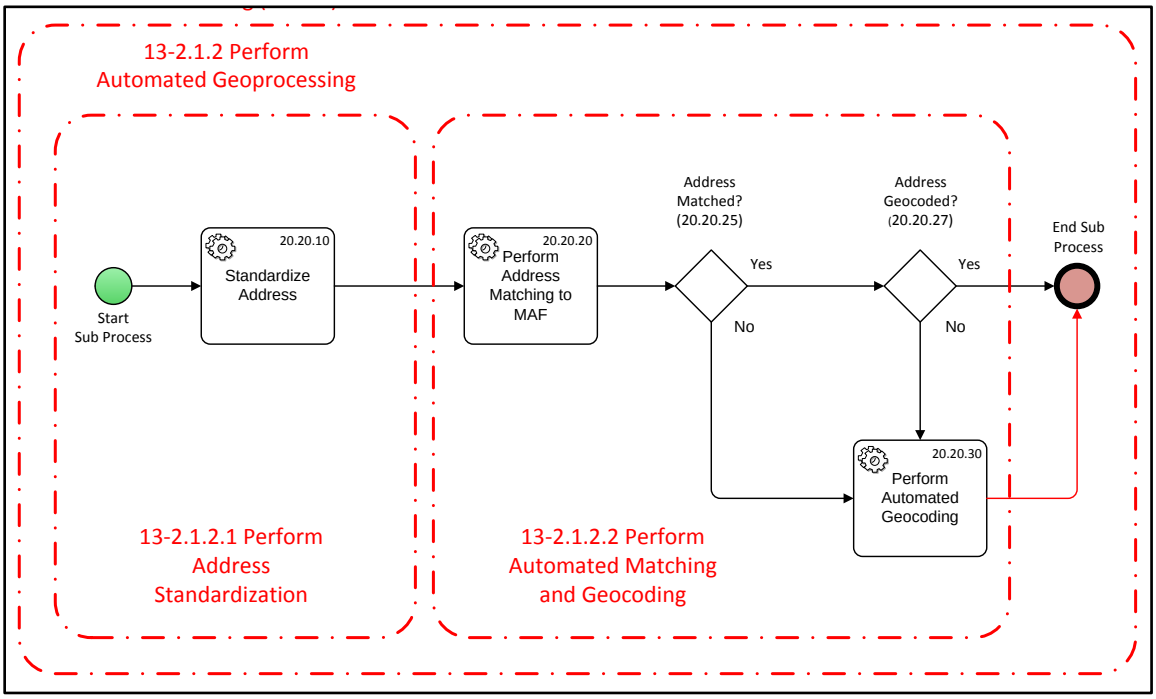

Figure 10: Perform Automated Geoprocessing

Perform Automated Geoprocessing is subdivided into the following Activity Areas.

Perform Address Standardization [13-2.1.2.1]

As shown in the BPM there is one step involved in this activity:

- Standardize Address [20.10.10]

Consolidated Non-ID cases received during Asynchronous Non-ID processing are standardized in preparation for matching and geocoding.

Perform Automated Matching and Geocoding [13-2.1.2.2]

As shown in the BPM, there are two steps involved in this activity:

- Perform Address Matching to MAF [20.20.20]
- Perform Automated Geocoding [20.20.30]

There is an attempt to match the standardized address to the MAF. If the address does not match a record in the MAF, then there is an attempt to geocode the address. If the address matches a record in the MAF but this record is not geocoded, there is an attempt to geocode the address.

3.2.1.3 Perform Unmatched Post Collection [13-2.1.3]

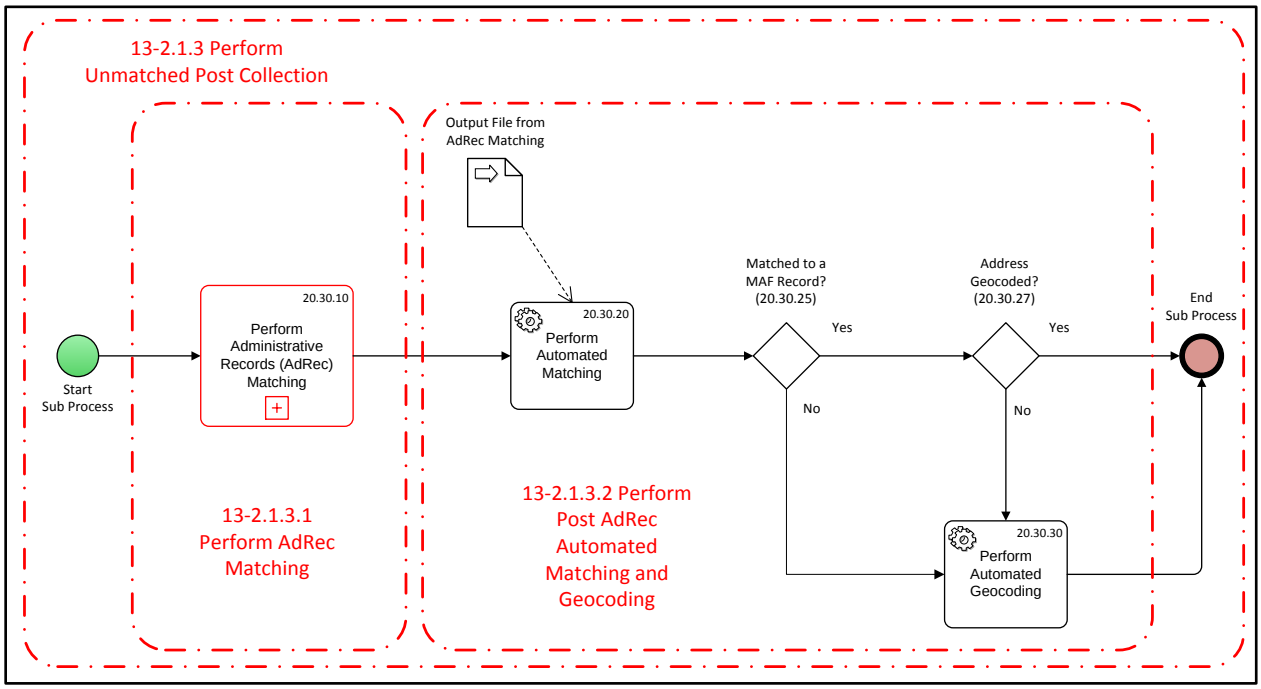

Figure 11: Perform Unmatched Post Collection

Perform Unmatched Post Collection is subdivided into the following Activity Areas.

- Perform Unmatched Post Collection [13-2.1.3]
  - Perform AdRec Matching [13-2.1.3.1]
  - Perform Post AdRec Automated Matching and Geocoding [13-2.1.3.2]

Perform AdRec Matching [13-2.1.3.1]

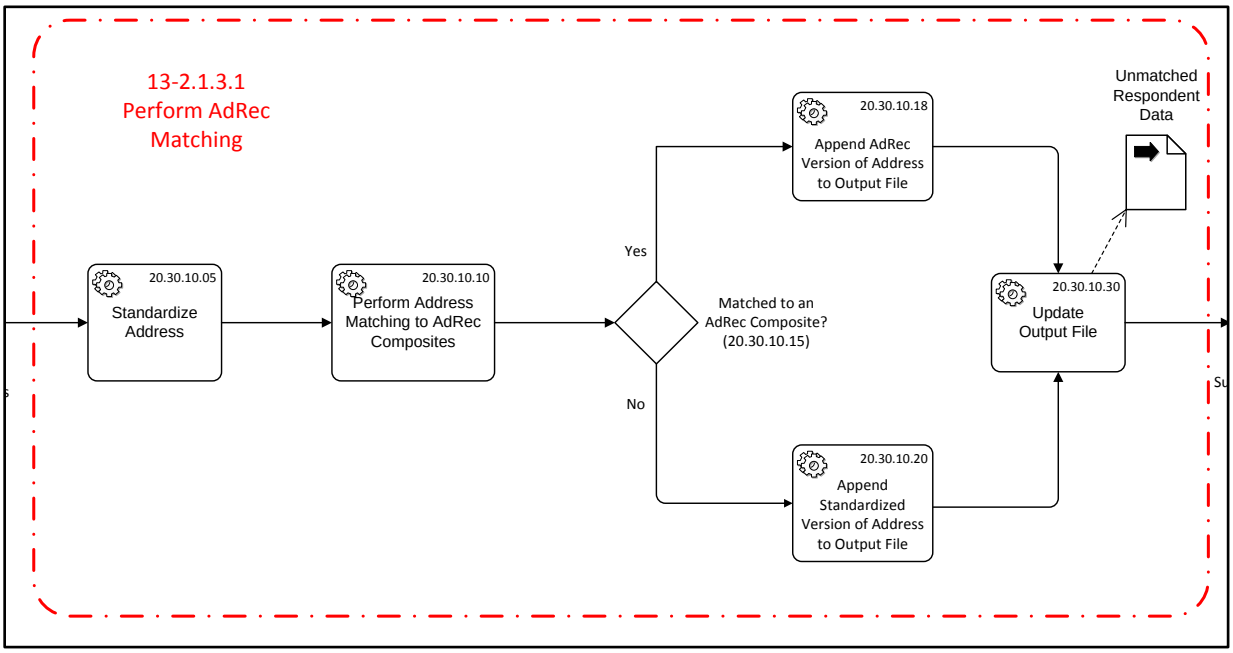

Figure 12: Perform AdRec Matching

As shown in the BPM, there are multiple steps involved in this activity:

- Standardize Address [20.30.10.05]
- Perform Address Matching to AdRec Composites [20.30.10.10]
- Append AdRec Version of Address to Output File [20.30.10.18]
- Append Standardized Version of Address to Output File [20.30.10.20]
- Update Output File [20.30.10.30]

Non-ID cases are standardized in preparation for AdRec matching. There is an attempt to match the standardized address to the AdRec composite file. If the address did not match the AdRec composite file, then the standardized version of the address is appended to the updated output file. If the address matches the AdRec composite file, then the AdRec version of the address is appended to the updated output file. Either the unmatched standardized or the matched AdRec address would now represent the best version of the address. This output file containing enhanced addresses will be put through automated matching and geocoding in subsequent activities.

Perform Post AdRec Automated Matching and Geocoding [13-2.1.3.2]

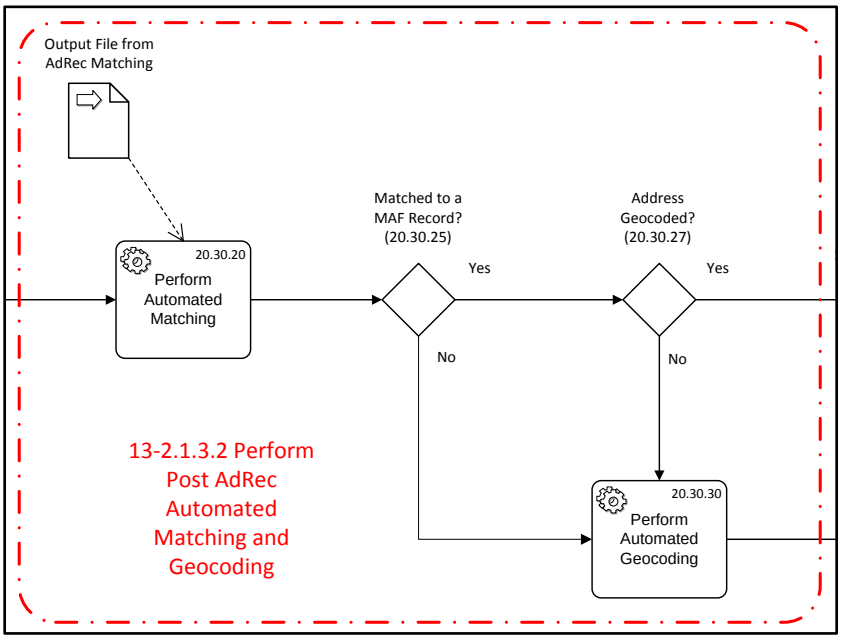

Figure 13: Perform Unmatched Post Collection

As shown in the BPM, there are potentially two steps involved in this activity:

- Perform Automated Matching [20.30.20]
- Perform Automated Geocoding [20.30.30]

There is an attempt to match the addresses on the output file from AdRec matching to the MAF. If an address did not match a record in the MAF, then there is an attempt to geocode that address. If an address matches a record in the MAF but this record was not geocoded, there is an attempt to geocode that address. Any address not matching to a MAF record is also sent on to the next phase: Manual Non-ID Processing.

3.2.2 Perform Manual Non-ID Processing [13-2.2]

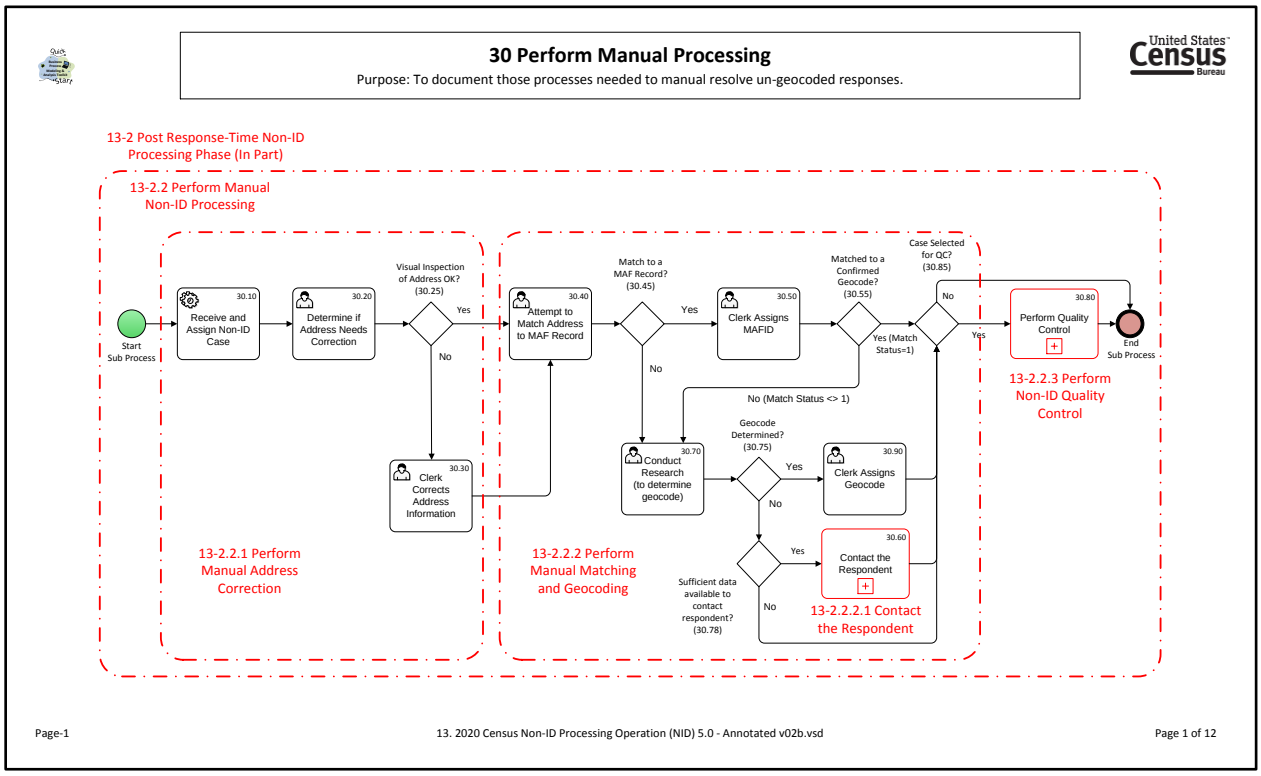

Figure 14: Perform Manual Non-ID Processing

Perform Manual Non-ID Processing is subdivided into the following Activity Areas.

- Perform Manual Address Correction [13-2.2.1]
- Perform Manual Matching and Geocoding [13-2.2.2]
    - Contact the Respondent [13-2.2.2.1]
- Perform Non-ID Quality Control [13-2.2.3]

3.2.2.1 Perform Manual Address Correction [13-2.2.1]

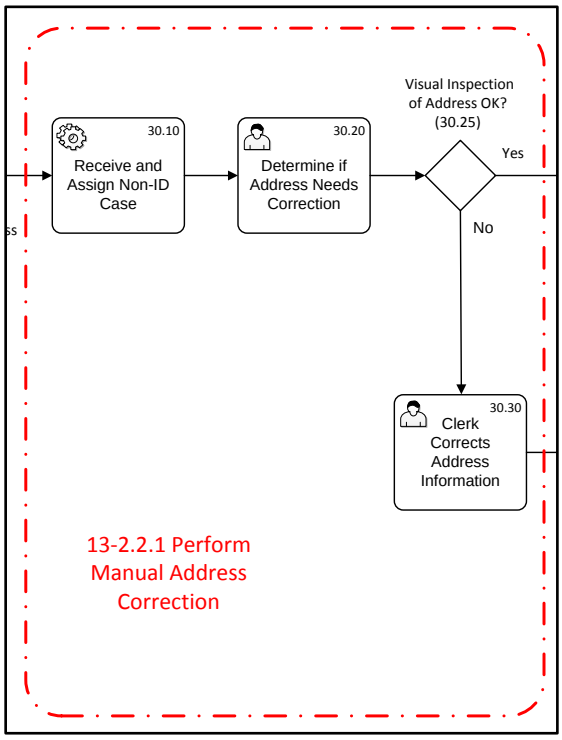

Figure 15: Perform Manual Address Correction

As shown in the BPM above, there are three steps involved in this activity:

- Receive and Assign Non-ID Case [30.10]
- Determine if Address Needs Correction [30.20]
- Clerk Corrects Address Information [30.30]

Perform Manual Address Correction is an attempt to correct or standardized address in preparation for matching and geocoding. This involves visually inspecting the address to determine if it needs correction. For example, a clerk might see a common misspelling of a street name, or something otherwise obvious to them but not detectable by an automated process. If the address does need correction, the clerk attempts to correct the address information.

3.2.2.2 Perform Manual Matching and Geocoding [13-2.2.2]

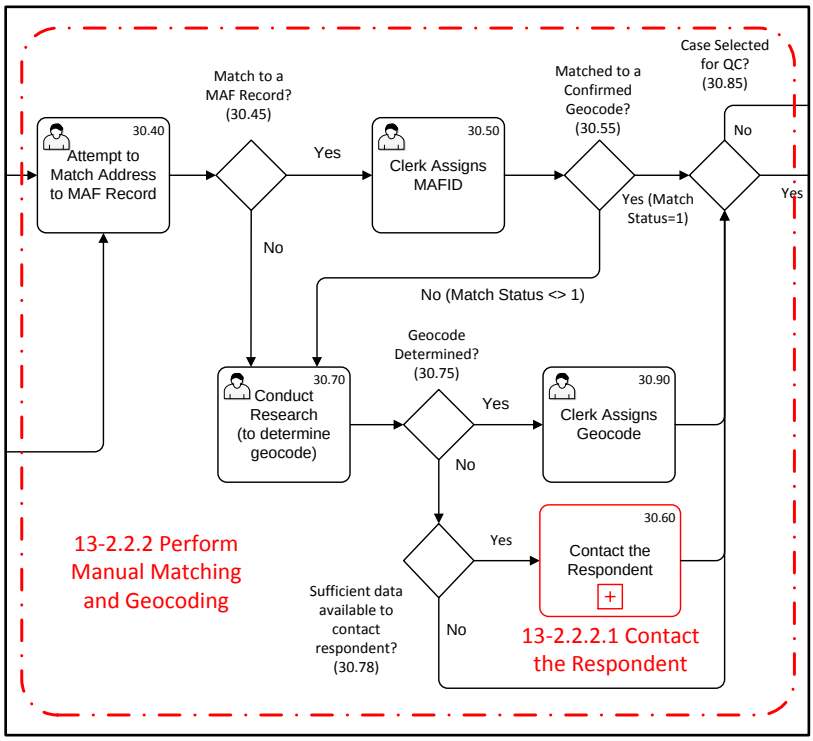

Figure 16: Perform Manual Matching and Geocoding

As shown in the BPM above, there are multiple steps involved in this activity:

- Attempt to Match Address to MAF Record [30.40]
- Clerk Assigns MAFID [30.50]
- Conduct Research (to determine geocode) [30.70]
- Clerk Assigns Geocode [30.90]
- Contact the Respondent [30.60]
    - Note: This business process is discussed in detail in section 13-2.2.2.1

Once a clerk accepts the corrected address there is an attempt to match the corrected address and/or geocode. First, the clerk attempts to match the corrected address to the MAF record and assign a MAFID. If the address does not match to a MAF record, then the clerk attempts to geocode the address. In order to geocode an unmatched MAF record, the clerk conducts research to determine the geocode. This includes contacting the respondent, if needed. In some cases, a respondent may even be able to provide sufficient further information to helps us match.

If the address is matched to a record in the MAF with a confirmed geocode, the matching and geocoding process is complete. If the address is not geocoded or is matched to a record in the

MAF that had an unconfirmed geocode, the clerk attempts to confirm the geocode. If the clerk is unable to locate the address or does not attempt to geocode, the matching and geocoding process is complete.

Contact the Respondent [13-2.2.2.1]

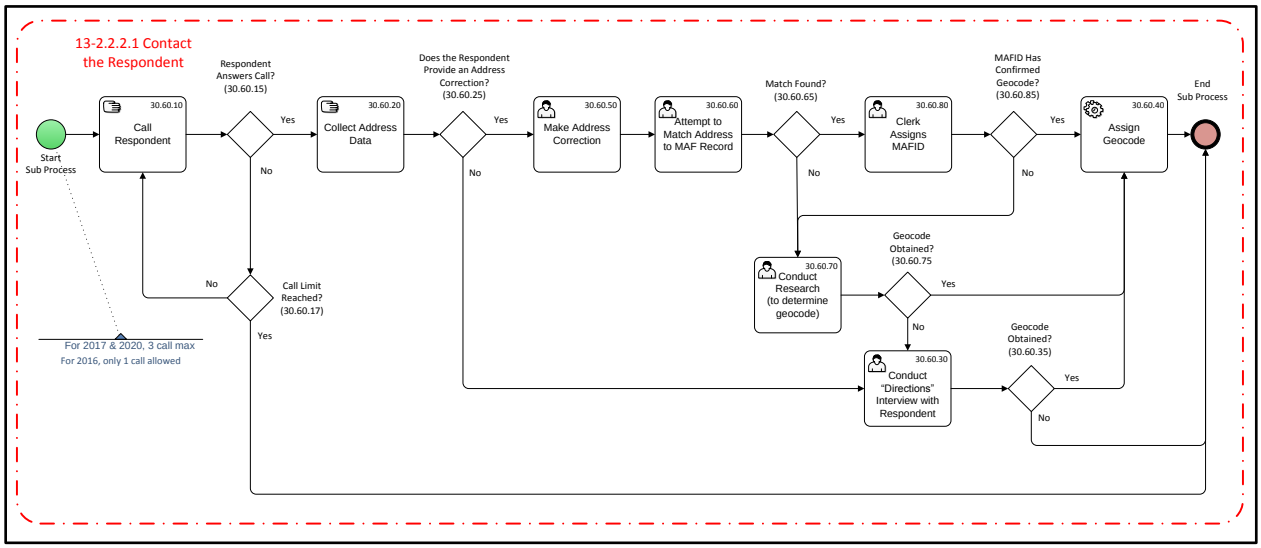

Figure 17: Contact the Respondent

As shown in the BPM, there are eight steps involved in this activity:

- Call Respondent [30.60.10]
- Collect Address Data [30.60.20]
- Conduct "Directions" Interview with Respondent [30.60.30]
- Make Address Correction [30.60.50]
- Attempt to Match Address to MAF Record [30.60.60]
- Clerk Assigns MAFID [30.60.80]
- Assign Geocode [30.60.40]
- Conduct Research (to determine geocode) [30.60.70]

Contact the Respondent includes calling the respondent to collect a corrected address and/or conducting a "Directions Interview" with the respondent to determine the geocode. If the respondent provides an address correction, the clerk updates the address with the corrected information and then tries to match the address to a MAF record. If the address does not match, then the clerk conducts research to determine the geocode.

3.2.2.3 Perform Non-ID Quality Control [13-2.2.3]

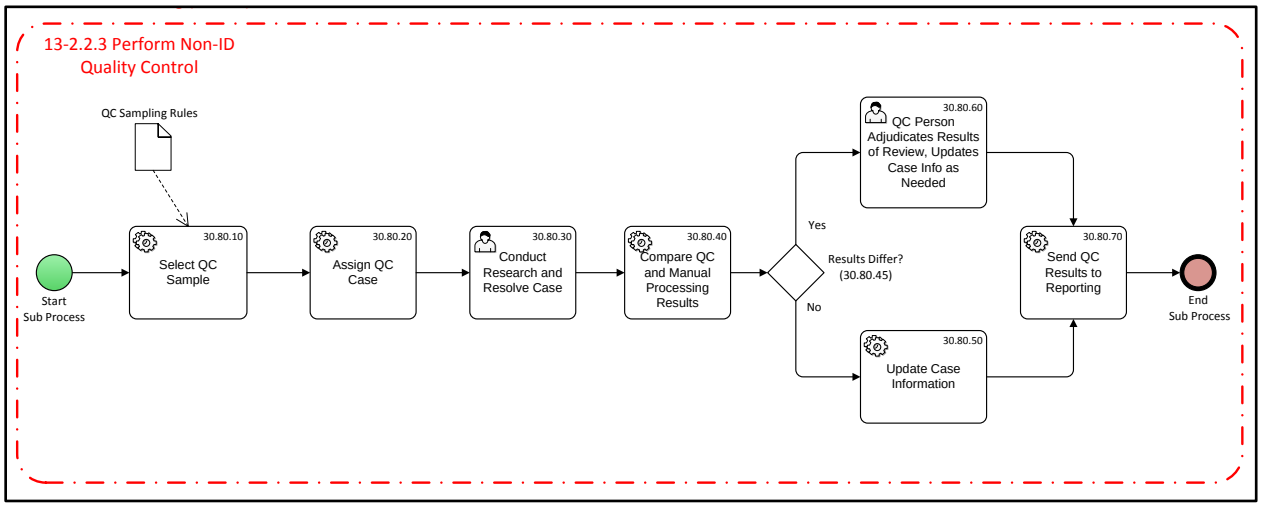

Figure 18: Perform Non-ID Quality Control

As shown in the BPM, there are seven steps involved in this activity:

- Select QC Sample [30.80.10]
- Assign QC Case [30.80.20]
- Conduct Research and Resolve Cases [30.80.30]
- Compare QC and Manual Processing Results [30.80.40]
- Update Case Information [30.80.50]
- QC Person Adjudicates Results of Review, Updates Case Information as Needed [30.80.60]
- Send QC Results to Reporting [30.80.70]

Perform Non-ID Quality Control includes comparing QC and Manual Processing results. Matching and Coding Software (MaCS) selects a QC sample and assigns a case to a QC reviewer. The QC reviewer then conducts research to resolve the case and compare the QC and manual processing results. If the results are different, the QC reviewer adjudicates the results of the review and updates case information as needed. If the results are not different, the QC reviewer updates the case information and sends the results to reporting.

3.2.3 Conduct Office-Based Non-ID Address Verification [13-2.3]

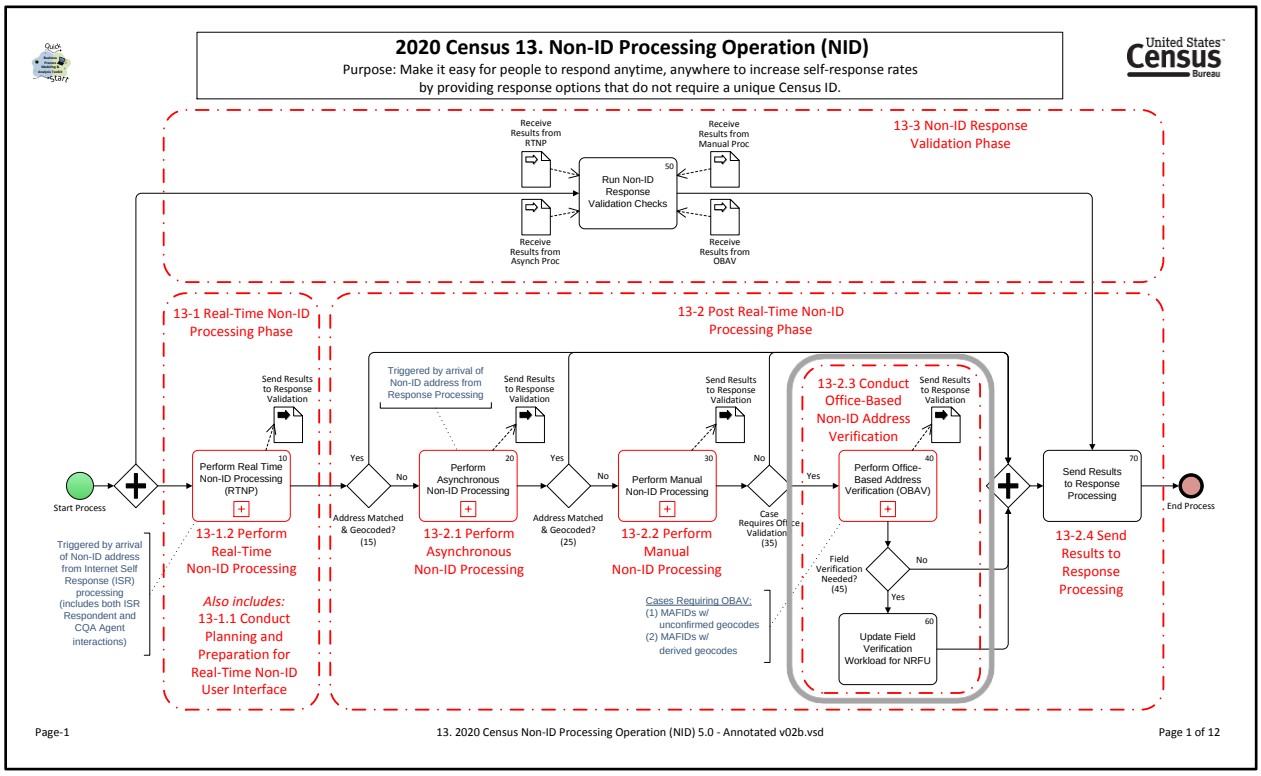

Figure 19: Conduct Office-Based Non-ID Address Verification

There are two steps involved in this activity:

- Perform Office-Based Address Verification (OBAV) [13-2.3.1]
- Update Field Verification Workload for NRFU [13-2.3.2]

Office-Based Non-ID Address Verification verifies the existence and census block location of LQs using geographic reference materials in an office-based operation. This work serves as the basis for longer-term efforts to reduce costs for the 2020 Census by avoiding fieldwork traditionally necessary to perform address confirmation for the Local Update of Census Addresses (LUCA) program and for eligible cases from Non-ID Processing.

3.2.3.1 Perform Office-Based Address Verification (OBAV) [13-2.3.1]

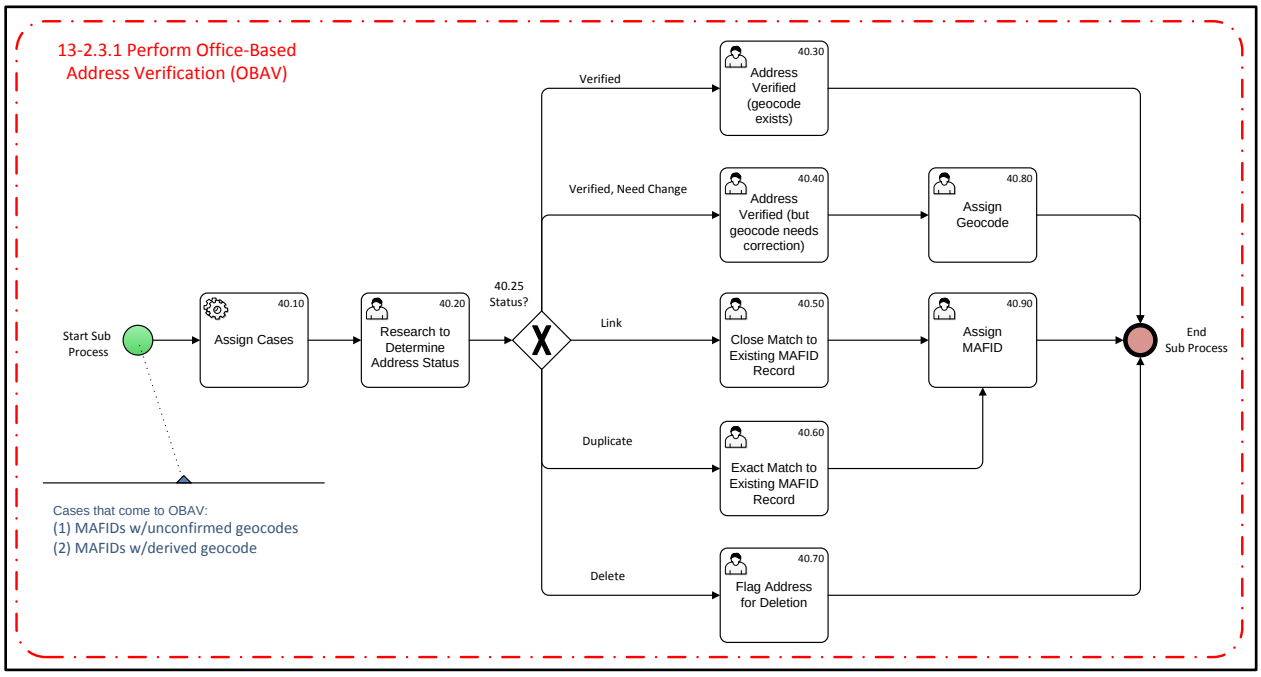

Figure 20: Perform Office-Based Address Verification (OBAV)

As shown in the BPM, there are nine steps involved in this activity

- Assign Cases [40.10]
- Research to Determine Address Status [40.20]
- Address Verified (geocode exists) [40.30]
- Address Verified (but geocode needs correction) [40.40]
- Close Match to Existing MAFID Record [40.50]
- Exact Match to Existing MAFID Record [40.60]
- Flag Address for Deletion [40.70]
- Assign Geocode [40.80]
- Assign MAFID [40.90]

Cases that come to OBAV are (1) MAFIDs with unconfirmed geocodes and (2) MAFIDs with derived geocodes. Clerks research to determine the address status. The outcome for these records will fall under four different main categories:

- Address and geocode were verified
- Address was verified but geocode was corrected
- Address was matched (linked) to an existing Census record

- Address could not be verified, and is therefore flagged for deletion (will not be included in Census data tabulation).

3.2.3.2 Update Field Verification Workload for NRFU [13-2.3.2]

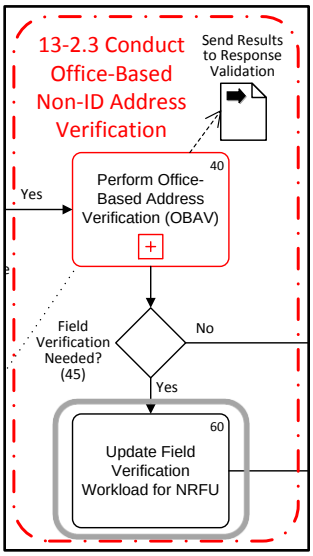

Figure 21: Update Field Verification Workload for NRFU

As shown in the BPM, there one step involved in this activity

- Update Field Verification Workload for NRFU [60]

Some cases from OBAV may require a field component to verify the location and existence of an address. These cases will make up the Field Verification workload that is sent to the Nonresponse Followup (NRFU) Operation.

3.2.4 Send Results to Response Processing [13-2.4]

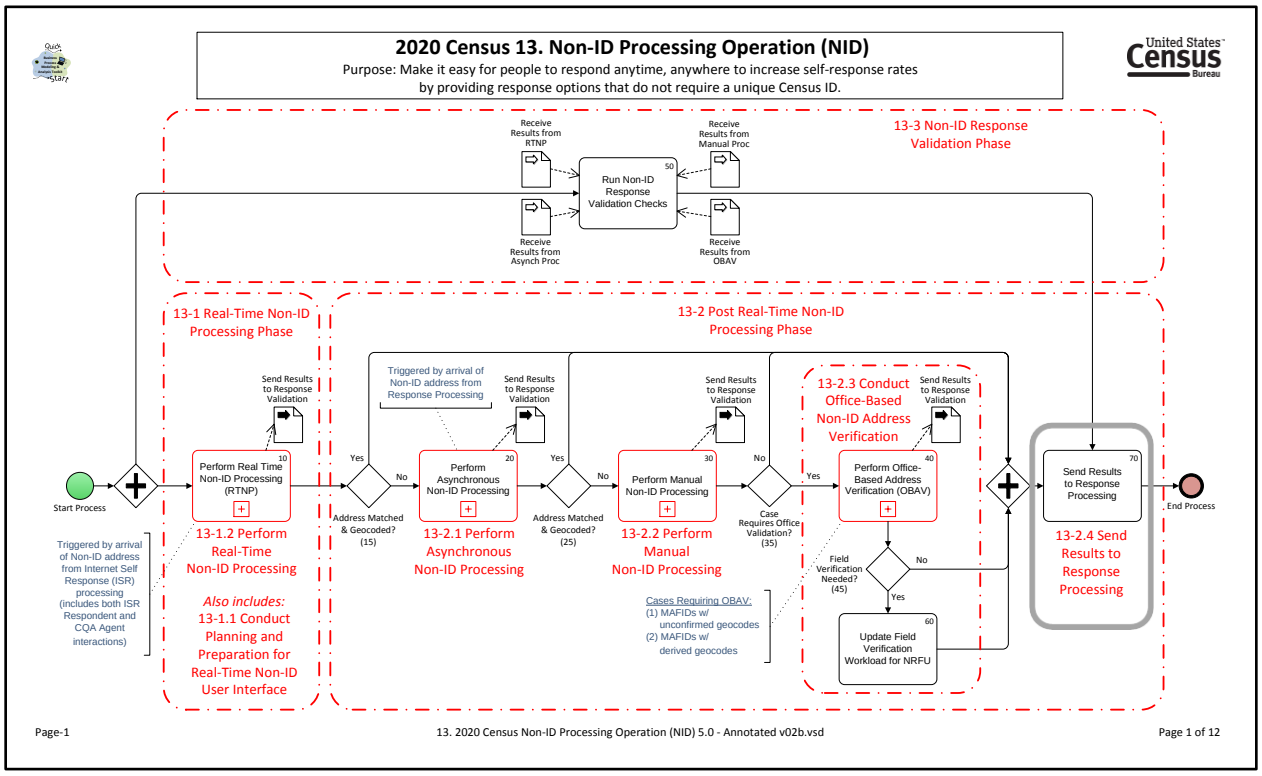

Figure 22: Send Results to Response Processing

There are two steps involved in this activity:

- Send Non-ID Results to RPO [13-2.4.1]
- Send Field Verification Workload to RPO for NRFU [13-2.4.2]

3.2.4.1 Send Non-ID Results to RPO [13-2.4.1]

Processing results from the three phases of NID (Real-Time, Post Real-Time, and Response Validation) are sent to RPO.

3.2.4.2 Send Field Verification Workload to RPO for NRFU [13-2.4.2]

Cases from OBAV that require a field component to verify the location and existence of an address and therefore have been included in the Field Verification workload are sent to NRFU through RPO.

3.3 Non-ID Response Validation Phase [13-3]

Figure 23 shows the BPM for the Non-ID Response Validation Phase [13-3] activity area (within the shaded gray rounded rectangle) and its constituent activities within the overall context of the NID Operation.

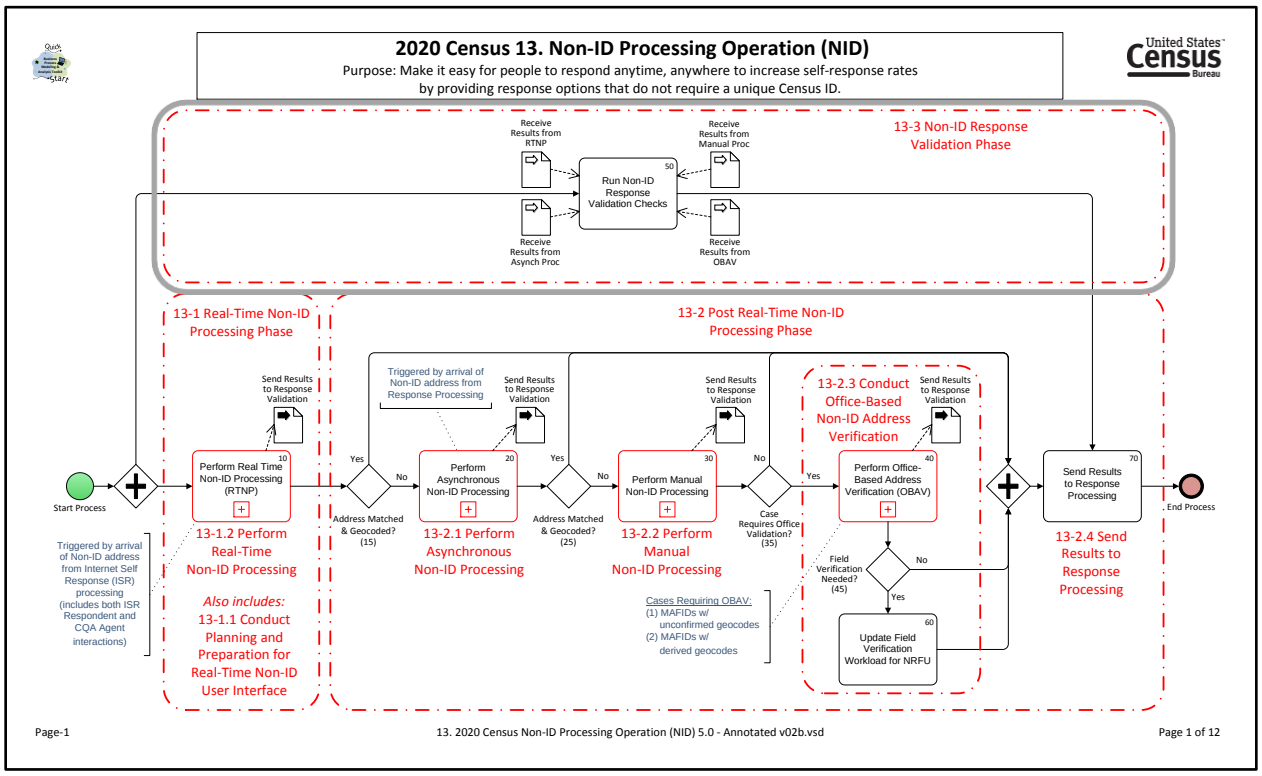

Figure 23: Non-ID Response Validation Phase [13-3] Constituent Activities

The Non-ID Response Validation Phase contains the following Activity.

- Non-ID Response Validation Phase [13-3]
    - Run Non-ID Response Validation Checks [13-3.1]

Subsequent sections describe the Non-ID Response Validation operational sub-activities in detail.

3.3.1 Run Non-ID Response Validation Checks [13-3.1]

Run Non-ID Response Validation Checks falls within the larger scope of the Fraud Detection process. The Run Non-ID Response Validation Checks box on the BPM indicates the point at which Non-ID responses will undergo the series of checks included in Fraud Detection.

The overall Fraud Detection process is depicted in the diagram below. Please note that the diagram is temporary, and will be replaced by a formal BPM once the Fraud Detection team has developed it.

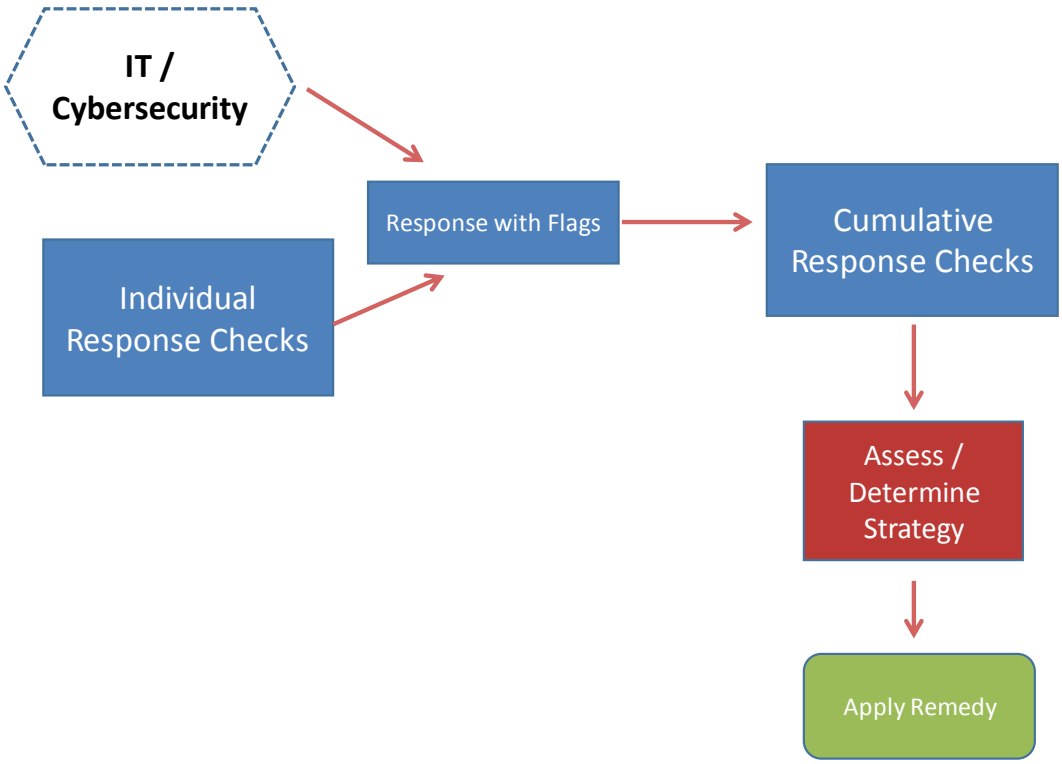

Figure 24: Overall Fraud Detection Workflow

In the overall fraud detection workflow, all responses (both ID and Non-ID) will undergo a series of checks at both an individual and cumulative level. These checks aim to detect potential individual fraud, as well as provide the data needed to detect potential widespread and targeted fraud by analyzing patterns found in the flagged responses. The details of this work are administratively restricted. However, a high-level description is provided below to explain the activities shown in Figure 24.

There will be a series of security measures built into the system used for Internet self-response in order to detect and address potential cyberattacks. Data from these monitoring efforts will be captured for analysis. In addition, every internet response will undergo a series of individual response checks. The result from this first stage of the validation process is an output file containing a series of flags depicting the results of the individual checks. At this stage, we will identify responses that exceed a certain threshold of suspicious indicators. Additionally, all the responses and their validation flags will be passed along to Response Processing for the next stage of the fraud detection flow.

Once Response Processing receives the output file containing the results of the validation checks conducted during responses, they will begin looking for larger suspicious patterns that may also indicate fraud. If any stage of the fraud detection process identifies suspicious behavior that

cannot be explained by reasonable sources, a determination will be made on whether any action must be undertaken, and if so, how to correct the fraud.

4. Cost Factors

4.1 Background

The activities associated with the planning and execution of the Non-ID Processing (NID) operation are not a significant cost for the 2020 Census; however, the investment in NID is projected to influence (reduce ↓ or increase ↑) the 2020 Census overall costs in the following ways:

- Increased self-response rates (↓)
- Improved coverage through self-response (↓)

4.2 Relevant IDEF0 Mechanisms

The following mechanisms from the IDEF0 Context Diagram represent the resources used to support this operation and comprise part of the 2020 Census cost elements:

Staff:

- HQ Staff
- NPC Staff

Sites:

- HQ
- NPC
- PRAO (TBD)

Systems:

- MAF/TIGER
- RTNP
- PEARSIS

Other Technology Infrastructure:

- HQ and NPC Office IT Infrastructure to conduct CFD operational work
- Census Network connectivity to transmit data between operational systems and operational sites

5. Measures of Success

For the 2020 Census operations, the corresponding Measures of Success will be documented in the operational assessment study plans and final reports. The operational assessment study plan documents the criteria that will be used to define successful completion of the operation. The operational assessment report will provide results on whether the criteria were met.

In general, operational assessments report on planned to actual variances in budget, schedules, and production and training workloads. The corresponding Measures of Success (as documented in the operational assessment study plan) include variances that exceed established thresholds. See *Content Guidelines for the 2020 Census Operational Assessments* for the potential scope of assessment.

Types of success measures include:

- **Process Measures** that indicate how well the process works, typically including measures related to completion dates, rates, and productivity rates
- **Cost Measures** that drive the cost of the operation and comparisons of actual costs to planned budgets. Costs can include workload as well as different types of resource costs
- **Measures of the Quality** of the results of the operation, typically including things such as rework rates, error rates, and coverage rates

See the corresponding operational assessment study plan and report for the 2020 Census Non-ID Processing Operation for details on the measures of success.

Appendix A – Acronyms and Terminology

The full list of 2020 Census acronyms and abbreviations are found in the List of Business Terms found in the 2020 Census Operational Plan.

Table 7 lists the specific acronyms and abbreviations used within this Detailed Operational Plan.

Table 8 lists a Glossary of Terms used within this Detailed Operational Plan.

Table 7: Acronyms and Abbreviations List

Acronym	Meaning
BPM	Business Process Model
BPMN	Business Process Model and Notation
CQA	Census Questionnaire Assistance
HQ	Headquarters
ID	Identification
IE	Information Exchanges
ISR	Internet Self Response
LQ	Living Quarters
MaCS	Matching and Coding Software
MAF	Master Address File
MTdb	MAF/TIGER Database
NID	Non-ID Processing Operation
NPC	National Processing Center
NRFU	Nonresponse Followup
OBAV	Office-Based Address Verification
PEARSIS	Production Environment for Administrative Records Staging, Integration, and Storage

Acronym	Meaning
QC	Quality Control
RPO	Response Processing Operations
RTNP	Real-Time Non-ID Processing
TIGER	Topologically Integrated Geographic Encoding and Referencing database

Table 8: Glossary of Terms

Term	Meaning
Asynchronous Processing	Processing performed post address collection from respondent
Clerical Processing	Interactive matching performed post address collection from respondent by staff
Non-ID Cases	Cases that do not have a preassigned Census ID
Real-Time Processing	Processing performed during address collection from respondent.

Appendix B – References

Appendix B lists the documents or other resources referenced within this Detailed Operational Plan document.

U.S. Census Bureau (2015), "2020 Census Operational Plan," Version 1.1, November 1, 2015.

U.S. Census Bureau (2016), "Operational Assessment Content Guidelines for the 2018 End-to-End Census Test and the 2020 Census," Draft, May 10, 2016.

Appendix C – Activity Tree for Non-ID Processing Operation (NID)

This appendix presents the Activity Tree for the NID Operation. An Activity Tree uses an outline structure to reflect the decomposition of the major operational activities in the operation. Each activity is numbered according to its position in the outline. For example, for the current operation numbered "13", the first activity would be numbered 13-1. Sub-activities under this activity would be numbered sequentially, starting again with the number one. For example, the first sub-activity under the first activity would be numbered 13-1.1 the second sub-activity as 13-1.2. The second activity would be numbered 13-2, and so on.

NID Activity Tree:

- 13-1 Real-Time Non-ID Processing Phase
  - 13-1.1 Conduct Planning and Preparation for Real-Time Non-ID User Interface
  - 13-1.2 Perform Real-Time Non-ID Processing
    - 13-1.2.1 Perform Real-Time Address Standardization
    - 13-1.2.2 Perform Real-Time Non-ID Matching and Geocoding
- 13-2 Post Real-Time Non-ID Processing Phase
  - 13-2.1 Perform Asynchronous Non-ID Processing
    - 13-2.1.1 Receive Consolidated Non-ID Cases
    - 13-2.1.2 Perform Automated Geoprocessing
      - 13-2.1.2.1 Perform Address Standardization
      - 13-2.1.2.2 Perform Automated Matching and Geocoding
    - 13-2.1.3 Perform Unmatched Post Collection
      - 13-2.1.3.1 Perform AdRec Matching
      - 13-2.1.3.2 Perform Post AdRec Automated Matching and Geocoding
  - 13-2.2 Perform Manual Non-ID Processing
    - 13-2.2.1 Perform Manual Address Correction
    - 13-2.2.2 Perform Manual Matching and Geocoding
      - 13-2.2.2.1 Contact the Respondent
    - 13-2.2.3 Perform Non-ID Quality Control
  - 13-2.3 Conduct Office-Based Non-ID Address Verification
    - 13-2.3.1 Perform Office-Based Address Verification (OBAV)
    - 13-2.3.2 Update Field Verification Workload for NRFU
  - 13-2.4 Send Results to Response Processing
    - 13-2.4.1 Send Non-ID Results to RPO
    - 13-2.4.2 Send Field Verification Workload to RPO for NRFU
- 13-3 Non-ID Response Validation Phase
  - 13-3.1 Run Non-ID Response Validation Checks

Appendix D – Business Process Models

This appendix includes all of the Annotated 2020 Census NID Operation Business Process Models (BPMs) for the NID Operation. The first sheet describes how to read the notation. Please refer to the *[TBD]* documents for additional information on the NID operational process flows.

NID Annotated BPM Diagram Set

PDF Attachment: Visio-13. 2020 Census Non-ID Processing Operation (NID) 5.0 - Annotated v02b.pdf

Visio-13. 2020 Census Non-ID Proces